

## **Passive Acoustic Monitoring for the Detection and Identification of Marine Mammals**

Marie A. Roch  
San Diego State University  
Department of Computer Science  
5500 Campanile Drive  
San Diego, CA 92182-7720  
phone: (619) 594-5830 fax: (619) 594-6746 e-mail: [marie.roch@sdsu.edu](mailto:marie.roch@sdsu.edu)

Award Number: N000140811199

### **LONG-TERM GOALS**

This project is intended to advance the state of passive acoustic monitoring of marine mammals. Improved methods of identifying cetaceans are developed in order to contribute to the Navy's mitigation efforts.

### **APPROACH**

This project is a multi-pronged study to advance the state of the field in three areas. The development of automated auditory scene analysis for delphinid tonal calls will permit subsequent work by this investigator or others to exploit the use of whistles for classification and localization. Our approach is to dynamically build hypothesis graphs to represent overlapping whistles. The whistle graphs are disambiguated using information from both sides of the crossings. In parallel to this effort, two modeling techniques are being pursued to improve existing passive acoustic monitoring capabilities based on echolocation clicks of odontocetes. The first of these examines the use of a universal background model as proposed by Reynolds et al. (2000) for human speaker verification tasks. Reynolds' problem, which is similar in nature to ours, is how can one reject vocalizations for which there is no data to create a model. We adapt his concept of a universal background model by training a generalized odontocete model using data from a number of species. Using Bayesian learning, training data from a specific species adapts the parameters of the generalized model. By having elements of the generalized model in both the adapted and background models, elements from an unknown species are more likely to be a better match for the general model, and those from the targeted species that differ from the training data are less penalized. The second approach for echolocation clicks exploits recent machine learning work on submanifold learning (Dasgupta and Freund 2007; Freund et al. 2007; Dasgupta and Freund 2008). In order to detect and classify odontocetes, features, or poignant characteristics of their signals, must be extracted from the audio signal. As the underlying process of sound generation cannot be measured directly, nor is it well understood, classification techniques must attempt to infer information about the producer of the signal (e.g. species) through a typically higher-order set of features. Submanifold learners focus on learning a subspace of the high-order feature space that can be more conducive to providing robust classification.

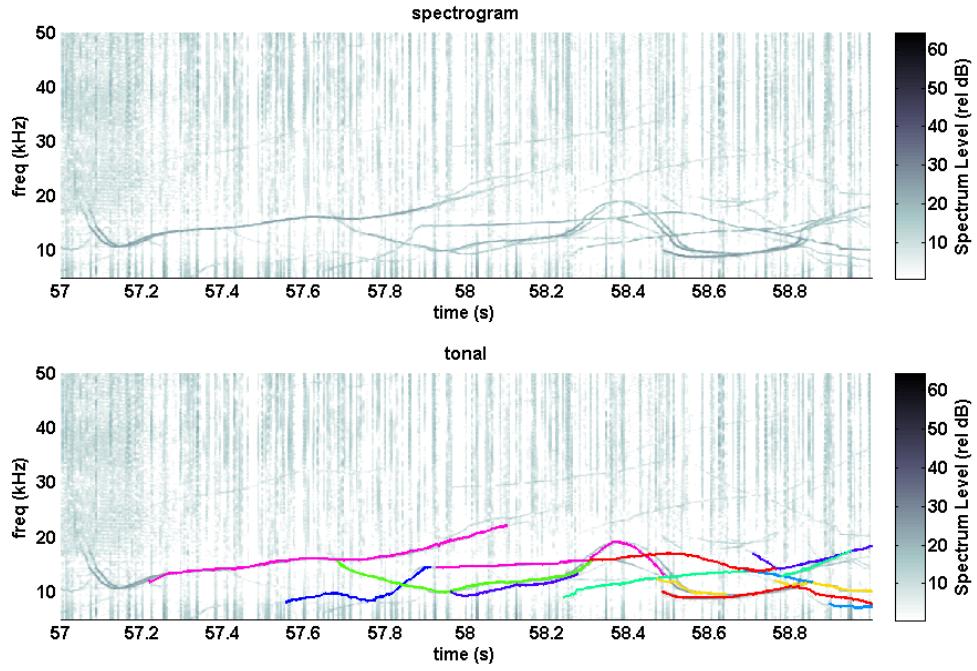
<b>Report Documentation Page</b>			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE <b>30 SEP 2011</b>	2. REPORT TYPE	3. DATES COVERED <b>00-00-2011 to 00-00-2011</b>		
<b>4. TITLE AND SUBTITLE</b> <b>Passive Acoustic Monitoring for the Detection and Identification of Marine Mammals</b>			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
<b>6. AUTHOR(S)</b>			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> <b>San Diego State University, Department of Computer Science, 5500 Campanile Drive, San Diego, CA, 92182-7720</b>			8. PERFORMING ORGANIZATION REPORT NUMBER	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> <b>Approved for public release; distribution unlimited</b>				
<b>13. SUPPLEMENTARY NOTES</b>				
<b>14. ABSTRACT</b>				
<b>15. SUBJECT TERMS</b>				
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b> <b>Same as Report (SAR)</b>	<b>18. NUMBER OF PAGES</b> <b>7</b>
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>		

## WORK COMPLETED

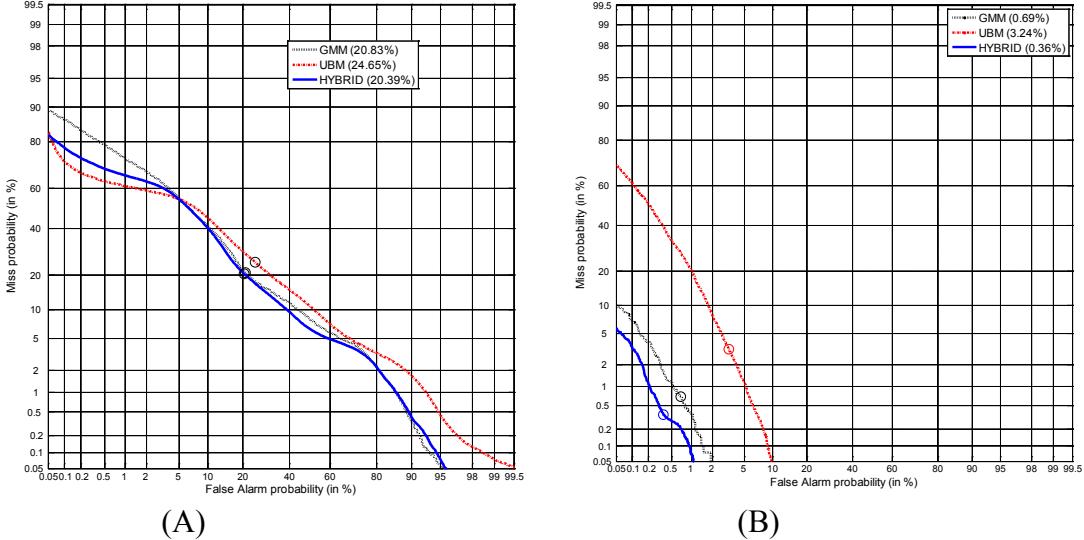
The whistle extraction system and universal background model detection system have been completed and evaluated on multi-species data set. We have created a framework for using the random projection tree submanifold learner and have extended an implementation of the algorithm provided by Freund and Dasgupta to provide pruning capabilities, a necessary component for tree-based classifiers which have a tendency to over learn when not pruned. An additional experiment using adaptive boosting (Freund and Schapire 1999) was conducted to provide additional comparison to the baseline.

## RESULTS

Automated methods for the detection of odontocete whistles have been developed and evaluated. Even in complex auditory scenes such as the long-beaked common dolphin calls shown in Figure 1, the system is capable of retrieving significant numbers of whistles. Determining the accuracy of this method presented non-trivial challenges. Techniques were developed for determining when a tonal which has a varying signal to noise ratio should be expected to be detected, as well as quantitative metrics of the detection quality such as percent of the tonal detected, percentage of the tonal detected, number of segments etc. We developed a tool to reduce the time required for human annotation of these whistles and developed a ground truth data set for recordings from five species in the Southern California Bight and Palmyra Atoll. We collaborated with Shannon Rankin (NOAA/NMFS) and Yvonne Barkley (Biowaves) to annotate a significant subset of our data, and this work is being carried on by Dave Mellinger's lab. These data and the methods for evaluating performance will be made publicly available as the classification task for the 2011 Detection, Classification, and Localization of Marine Mammals Using Passive Acoustics conference. Formal evaluation was conducted on about an hour of annotated data from five different species: short- and long-beaked common dolphins, bottlenose dolphins, melon-headed whales, and spinner dolphins. On average, our system is capable of retrieving about 80% of the whistles meeting SNR and duration criteria. Approximately 85% of each whistle is retrieved with an average deviation of less than one frequency bin from the human detections. Of the detections, nearly 77% were detections of valid whistles, with the false positives being primarily attributable to regions with highly variable noise and echosounders. Our analysis has led to several planned optimizations that could significantly increase the percentage of valid detections in the future. A manuscript is in progress and will be submitted before the end of the performance period of this award.



**Figure 1 (color online)** - Whistles extracted from short beaked common dolphin (*Delphinus capensis*) whistles with a threshold of 10 dB relative sound pressure level. The spectrogram is shown with (lower) and without (upper) detections overlaid. Common dolphins aggregate in large groups and frequently have many overlapping whistles.



**Figure 2 (color online) – Detection error tradeoff curves (Martin et al. 1997) for a species detection task when examples of echolocation clicks from the impostor species has not been seen in the training data. Curves that are closer to the origin have better performance. A circle shows the point on each curve where the miss and false alarm probabilities are equal with the associated error rate listed in the legend. (a) overall performance of the baseline GMM system (black dotted), universal background model (red dash-dot), and the hybrid system (blue solid). (B) detection of presumed cuvier's beaked whale clicks.**

As reported last year, experiments with the universal background model suggested that the relatively small number of species available for the background model may have limited the effectiveness of this technique. As significantly increasing the number of species used in training was not a realistic option, we evaluated a hybrid model based on the concept of deleted interpolation (Huang et al. 2001). Deleted interpolation provides a method of fusing scores from two models through a simple weighting scheme. Typically, the expectation maximization algorithm is used to determine the weighting, but this would have required additional training data which was not available. A uniform prior was assumed, and the hybrid model resulted in modest improvements over baseline techniques (Figure 2A). For some species, such as data from presumed Cuvier's beaked whales (Figure 2B), the improvement in performance was more pronounced. Insights gained from these experiments led to changes in feature extraction in a related species classification task that resulted in over a 15% reduction in error rate in a species identification task on five Southern California odontocetes as compared to our previous methods (Roch et al. 2008) on the same data. These findings have been accepted for publication (Roch et al. in press).

The final study sponsored by this work is our analysis of the submanifold learning algorithm RP-Tree. We modified Dasgupta and Freund's (2008) algorithm to create classification trees and compensated for overtraining using Quinlan's pruning rules (1993). We compared the submanifold learner with our previous work using Gaussian mixture models and the adaptive boosting algorithm (Hastie et al. 2009),

an algorithm that is widely respected in the machine learning community but had not yet been applied to the classification of marine mammal sounds. This permitted us to compare both generative and discriminative models (the GMMs and adaptive boosting respectively) with a submanifold learner. The submanifold learner was unable to outperform either learning algorithm, and had an approximately 20% higher error rate than the best performing algorithm. Adaptive boosting performance was roughly equivalent to GMMs. While the results of the study did not demonstrate an ability to find effective submanifolds in our feature set, it has engendered several new ideas for feature extraction for which we are pursuing preliminary experiments.

In summary, this project has reduced error rates on a best in class recognition system by 15% (Roch et al. in press), and produced a whistle extraction algorithm (Roch et al. in manuscript) capable of analyzing multiple overlapping whistles in very complex auditory scenes involving large groups of concurrently whistling dolphins.

## **IMPACT/APPLICATIONS**

Methods developed in this research for echolocation click classification have been rewritten as production code and are being developed as part of a system to provide real-time detection of species for autonomous underwater gliders (N00014-08-1124). Dave Moretti's group (NUWC) has expressed interest in running these classifiers on the SCORE and AUCTEC ranges and we are working with his group to make this a reality.

Our whistle annotation data will be made available to groups developing whistle classification algorithms as part of the 2011 Detection, Classification, and Localization of Marine Mammals Using Passive Acoustics conference. The automated extraction of whistles paves the way for the development of new classification algorithms that exploit information about whistles beyond simple contour statistics such as minimum, maximum frequency, number of inflection points, etc. Members of the Hildebrand lab have also been conducting preliminary localization experiments by beam-forming the automated whistle detections. Early experiments show the potential to track group fusion/fission behavior.

## **RELATED PROJECTS**

Project title: Southern California Marine Mammal Studies; John Hildebrand, Principal Investigator.

Sponsor: CNO N45 and the Naval Postgraduate School; Support from this project allowed for collection of the acoustic data used to create the southern California marine mammal call database.

Project title: Glider-based Passive Acoustic Monitoring Techniques in the Southern California Region. ONR Grant: N000140811124. This project used algorithms from this project for marine mammal detection and classification.

Project title: Flying wing underwater glider for persistent surveillance missions, Gerald L. D'Spain Principal Investigator. ONR Grant: N000140410558. This project supported development of large autonomous underwater gliders based on the flying wing design.

## REFERENCES

- Dasgupta, S., Y. Freund (2007) Random projection trees and low dimensional manifolds. *Dept. of Computer Science and Engineering*, University of California, San Diego, San Diego. CS2007-0890, 29.
- Dasgupta, S., Y. Freund (2008) Random projection trees and low dimensional manifolds, Symp. on Theory of Comp. (STOC), Victoria (BC), Canada: 10.
- Freund, Y., S. Dasgupta, M. Kabra, N. Verma (2007) Learning the structure of manifolds using random projections, Adv. in Neural Info. Proc. Systems (NIPS), Vancouver, B.C., Canada.
- Freund, Y., R.E. Schapire (1999) A Short Introduction to Boosting (Japanese, English version available at <http://cseweb.ucsd.edu/~yfreund>). *J. Japanese Soc. Artif. Int.*, **14**(5), 771-780.
- Hastie, T., R. Tibshirani, J.H. Friedman (2009). *The elements of statistical learning data mining, inference, and prediction*. Springer, New York: xxii, 745 p. pp.
- Huang, X., A. Acero, H.W. Hon (2001). *Spoken Language Processing*. Prentice Hall PTR, Upper Saddle River, NJ.
- Martin, A., G. Doddington, T. Kamm, M. Ordowski, M. Przybocki (1997) The DET curve in assessment of detection task performance, Eurospeech, Rhodes, Greece: 1895-1898.
- Quinlan, J.R. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann, San Mateo.
- Reynolds, D.A., T.F. Quatieri, R.B. Dunn (2000) Speaker verification using adapted Gaussian mixture models. *Digital Signal Processing*: 19-41.
- Roch, M.A., S. Brandes, B. Patel, Y. Barkley, S.R. Rankin, S. Baumann-Pickering, M.S. Soldevilla, J.A. Hildebrand (**in manuscript**) Automated extraction of odontocete whistle contours.
- Roch, M.A., H. Klinck, S. Baumann-Pickering, D.K. Mellinger, S. Qui, M.S. Soldevilla, J.A. Hildebrand (**in press**) Classification of echolocation clicks from odontocetes in the Southern California Bight. *J. Acous. Soc. Am.*
- Roch, M.A., M.S. Soldevilla, R. Hoenigman, S.M. Wiggins, J.A. Hildebrand (2008) Comparison of machine learning techniques for the classification of echolocation clicks from three species of odontocetes. *Canadian Acoust.*, **36**(1), 41-47.

## PUBLICATIONS AND PRESENTATIONS

- Roch, M.A., H. Klinck, S. Baumann-Pickering, D.K. Mellinger, S. Qui, M.S. Soldevilla, J.A. Hildebrand (**in press**) Classification of echolocation clicks from odontocetes in the Southern California Bight. *J. Acous. Soc. Am.*

M. A. Roch, B. Patel, S. Rankin, Y. Barkley, M. S. Soldevilla, J. A. Hildebrand, “Identifying delphinid whistles contours using graph search,” invited presentation Acoust. Soc. Am., Baltimore, MD, April 2010.

S. Qui, M. A. Roch, M. S. Soldevilla, and J. A. Hildebrand, “Rejecting unknown species in a species detection task,” Intl. Workshop on the Detection and Classification of Marine Mammals Using Passive Acoustics, Pavia, Italy, September 2009.

M. A. Roch, B. Patel, M. S. Soldevilla, and J. A. Hildebrand, “Identifying delphinid contours using graph search,” Intl. Workshop on the Detection and Classification of Marine Mammals Using Passive Acoustics, Pavia, Italy, September 2009.